GRAŻYNA DEHNEL
JAN PARADYSZ

WYBRANE ESTYMATORY UWZGLĘDNIAJĄCE INFORMACJE O DODATKOWEJ CESZE A STATYSTYKA MAŁYCH OBSZARÓW

Wykorzystanie informacji o parametrach rozkładu dodatkowej cechy X przy estymacji pewnych parametrów cechy Y było przedstawiane w polskiej literaturze przedmiotu, por. np. C. Bracha [1987], C. Bracha, B. Lednicki, [1989], R. Zasępa [1972], J. Wywiał [1992]. Jednakże była to relacja raczej o osiągnięciach za granicą niż o polskich doświadczeniach badawczych. Ponadto uwzględniono jedynie przypadki klasyczne estymacji ilorazowych i regresyjnych, pomijając dynamicznie rozwijającą się w ostatnim ćwierćwieczu statystykę małych obszarów. Powrót Polski do krajów o racjonalnej gospodarce rynkowej pozwala przypuszczać, że będzie teraz duże zapotrzebowanie społeczne na metody uwzględniające w większym stopniu koszty badań. Jednym z pierwszych objawów takiej tendencji była konferencja zorganizowana w Warszawie w dniach 30 września – 3 października 1992 r. przez Główny Urząd Statystyczny i Międzynarodowy Instytut Statystyczny.

Celem niniejszego opracowania jest:

– zwrócenie uwagi, że oprócz tradycyjnych estymatorów ilorazowych i regresyjnych istnieją także możliwości stosowania estymatorów opartych na innych formułach (np. estymatory różnicowe, estymatory oparte na sumie),

– poszerzenie listy estymatorów o nowe propozycje (estymatory ilorazu, różnicy i sumy),

– ukazanie możliwości zastosowań statystyki małych obszarów w Polsce.

1. Szacowanie średniej ($\bar{Y}$) z uwzględnieniem informacji o dodatkowej cesze X

Szacując $\bar{Y}$ na podstawie próby losowej, można znacznie poprawić efektywność estymatora, korzystając z informacji odnośnie do parametrów rozkładu dodatkowej cechy Y bądź dwuwymiarowego rozkładu tych cech.

Najczęściej stosowanymi estymatorami są (por. Zasępa [1972, s. 76-98]):

a) ilorazowy,

$$\bar{y}_q = \bar{y} \cdot \frac{\bar{X}}{\bar{x}} = \frac{\bar{y}}{\bar{x}} \cdot \bar{X} = r \cdot \bar{X}, \tag{1}$$

b) liniowo-regresyjny,

$$\bar{y}_{lr} = \bar{y} + b(\bar{X} - \bar{x}), \tag{2}$$

gdzie: $\bar{x}$, $\bar{y}$ – średnie z próby prostej,
$\bar{X}$ – znana średnia z populacji generalnej,
b – współczynnik równania regresji cechy y względem cechy x w próbie.

Wariancje estymatorów (1) i (2) są odpowiednio równe

$$D^2(\bar{y}_q) = \left(1 - \frac{n}{N}\right)\left[\frac{R^2 S_x^2 + S_x^2 + S_y^2 - 2\rho_{xy} R S_x S_y}{n}\right], \tag{3}$$

$$D^2(\bar{y}_{lr}) = \left(1 - \frac{n}{N}\right)\frac{S_y^2(1 - \rho_{xy}^2)}{n}, \tag{4}$$

gdzie odnoszące się do populacji generalnej parametry oznaczają:
S_x^2 – wariancja cechy X,
S_y^2 – wariancja cechy Y,
ρ_{xy} – współczynnik korelacji liniowej między X i Y,

$$R = \frac{\bar{Y}}{\bar{X}}.$$

Obydwa estymatory są zgodne i obciążone. Obciążenie estymatora (1) maleje wraz ze wzrostem liczebności próby (n), frakcji, jaką stanowi próba w populacji generalnej $\left(\frac{n}{N}\right)$ oraz wzrostu współczynnika korelacji między cechami.

$$\text{bias}\,(\bar{y}_q) \approx \left(1 - \frac{n}{N}\right)\frac{R S_x^2 - \rho_{xy} S_x S_y}{n\bar{X}}. \tag{5}$$

Obciążenie estymatora regresyjno-liniowego można przedstawić za pomocą następującej formuły:

$$\text{bias}\,(\bar{y}_{lr}) \approx \left(1 - \frac{n}{N}\right) \cdot \frac{1}{(n-1)S_x^2} \cdot \frac{1}{N-1} \cdot \left[B\sum(X_k - \bar{X})^3 - \sum(X_k - \bar{X})^2(Y_k - \bar{Y})\right]. \tag{6}$$

Obciążenie (6) zależy od liczebności próby, frakcji $\frac{n}{N}$ oraz postaci liniowej funkcji regresji. Jeśli regresja y od x jest liniowa, to wyrażenie (2) jest nieobciążonym estymatorem parametru $\bar{Y}$, por. W. G. Cochran [1976, s. 217].

W literaturze przedmiotu zaproponowano estymatory i ilorazowe, i regresyjne, które nie są obciążone. Do tego rodzaju estymatorów złożonych należą między innymi (por. C. Bracha [1987, s. 38]):

c) Estymator ilorazowy Hartleya-Rossa

$$\bar{y}_{qHR}=r\bar{X}+\frac{n(N-1)}{N(n-1)}\bar{y}-r\bar{x}, \tag{7}$$

gdzie

$$r=\frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i}. \tag{8}$$

d) Estymator M. R. Mickeya

$$\bar{y}_{qM}=\bar{r}_k\bar{X}+\frac{[(N-m)(k-1)]\cdot k}{N}(\bar{y}-\bar{r}_k\bar{x}), \tag{9}$$

$$\bar{r}_k=\frac{1}{k}\sum_{i=1}^{k}r_i, \qquad r=\bar{y}/\bar{x}, \tag{10}$$

gdzie: k – liczba podprób, na które została podzielona już po wylosowaniu próba,
m – liczba elementów w podpróbie,
i – numer podpróby,
$i=1,\ldots,k$,
$\bar{x}$, $\bar{y}$ – średnie dla całej próby.

Wśród nieobciążonych estymatorów regresyjnych C. Bracha wyróżnia:

e) Estymator M. R. Mickeya

$$\bar{y}_{lr\,M}=\bar{y}+\bar{b}'_k(\bar{X}-\bar{x})+\frac{N-n}{Nk}\left(\sum_{i=1}^{k}\bar{x}_i b'_i-n\bar{x}b'_k\right), \tag{11}$$

gdzie: $\bar{b}'=\frac{1}{k}\sum_{i=1}^{k}b_i$,

b'_i – współczynnik regresji cechy Y względem X otrzymany z próby, w której pominięto i-tą podpróbę, $i=1,\ldots,k$,

f) Estymator W. H. Williamsa

$$\bar{y}_{lr\,w}=\bar{y}+\bar{b}_k(\bar{X}-\bar{x})+\frac{N-n}{Nk(k-1)}\left(\sum_{i=1}^{k}\bar{x}_i b_i-n\bar{x}\bar{b}_k\right), \tag{12}$$

gdzie: $\bar{b}_k=\frac{1}{k}\sum_{i=1}^{k}b_i$,

b_i – współczynnik regresji cechy Y względem cechy X z i-tej podpróby, $i=1,\ldots,k$.

Obszerniej na temat estymatorów c)-f) pisze C. Bracha [1987], gdzie można także znaleźć informacje odnośnie do tworzenia podprób oraz technik estymacji.

Zdaniem C. Brachy [1987, s. 39], redukcja lub eliminacja obciążenia metodą podprób poprawia efektywność estymatorów ilorazowych w porównaniu z klasycznymi tylko przy bardzo małych liczebnościach (rzędu kilku jednostek). W przypadku liczniejszych prób tego rodzaju zabiegi nie mają praktycznego znaczenia.

Oprócz klasycznych estymatorów przedstawionych w podręcznikach (ilorazowy i regresyjny) możliwe są także estymatory oparte na innych formułach. Celem oszacowania $\bar{Y}$ można wykorzystać także estymatory: iloczynowy i różnicowy.

g) Estymator iloczynowy o postaci

$$\bar{y}_{pr}=\bar{y}\left(\frac{\bar{x}}{\bar{X}}\right)=\frac{1}{\bar{X}}\cdot w_p. \tag{13}$$

Estymator (13) jest estymatorem obciążonym, którego obciążenie maleje wraz ze wzrostem n. Jego wariancja wynosi w przybliżeniu

$$D^2(\bar{y}_{pr})\approx\frac{1}{\bar{X}^2}D^2(w_p)\approx\left(\frac{N-n}{Nn}\right)\left[S_y^2+\left(\frac{\bar{Y}}{\bar{X}}\right)^2 S_x+2\left(\frac{\bar{Y}}{\bar{X}}\right)\rho_{xy}S_xS_y\right]. \tag{14}$$

Jeśli wariancję estymatora ilorazowego przedstawić jako

$$D^2(\bar{y}_q)=\left(\frac{N-n}{Nn}\right)\left[S_y^2+\left(\frac{\bar{Y}}{\bar{X}}\right)S_x^2-2\rho_{xy}S_xS_y\left(\frac{\bar{Y}}{\bar{X}}\right)\right], \tag{15}$$

to widać wyraźnie, że estymator (13) jest efektywniejszy od estymatora ilorazowego wówczas, gdy zależność pomiędzy zmiennymi X i Y jest ujemna. Jeśli zależność jest dodatnia, to estymator ilorazowy jest efektywniejszy od estymatora iloczynowego. W przypadku braku współzależności ($\rho_{xy}=0$) oba estymatory są jednakowo efektywne, a od obydwóch znaczniej efektywniejszym estymatorem jest średnia z próby ($\bar{y}$).

h) Estymacja średniej $\bar{Y}$ na podstawie różnicy średnich cechy dodatkowej $(\bar{X}-\bar{x})$.

W jednym z wariantów estymatora regresyjnego zakłada się, że estymator współczynnika regresji b parametru B otrzymuje się nie na podstawie próby, na przykład za pomocą metody najmniejszych kwadratów, ale ustala się dowolnie

jako pewnego rodzaju stałą, powiedzmy b_0. W takim przypadku estymator parametru $\bar{Y}$ znajduje się analogicznie jak estymator regresyjny

$$\bar{y}_{diff} = \bar{y} + b_0(\bar{X} - \bar{x}). \tag{16}$$

Równanie (16) można napisać także w postaci

$$\bar{y}_{diff} = (\bar{y} - b_0\bar{x}) + b_0\bar{X}. \tag{17}$$

Ponieważ b_0 jest dowolną stałą, więc można ją tak wybrać, żeby otrzymać przypadki najprostsze, np. $b_0 = 1$, wówczas

$$\bar{y}_{diff} = (\bar{y} - \bar{x}) + \bar{X}, \tag{18}$$

gdzie estymator (18) jest zwykłą różnicą dwóch zmiennych losowych.

Jeśli $b_0 \neq 1$, to wyrażenie w nawiasie równanie (18) można traktować jak różnicę ważoną. Ten estymator jest estymatorem nieobciążonym, gdyż

$$E(\bar{y}_{diff}) = E[\bar{y} - b_0(\bar{x} - \bar{X})] = \bar{Y} - b_0\bar{x} - b_0\bar{X} = \bar{Y},$$

natomiast dyspersja estymatora (17)

$$D^2(\bar{y}_{diff}) = \left(\frac{N-n}{Nn}\right) \cdot (S_y^2 + b_0^2 S_x^2 - 2b_0\rho_{xy}S_xS_y). \tag{19}$$

Dokładność estymatora (17) podobnie jak estymatora ilorazowego oraz regresyjnego „zwykłego" zależy od korelacji między x i y oraz od wyboru b_0. Można wykazać, że najlepszym wyborem b_0 jest wybór spełniający relację (por. R. J. Jessen [1985, s. 129]):

$$b_0 = \frac{S_{xy}}{S_x^2} = B, \tag{20}$$

to znaczy b_0 oblicza się za pomocą metody najmniejszych kwadratów na realizacjach x i y.

2. Szacowanie sumy, różnicy, wskaźnika natężenia i dynamiki

W naukach społecznych i ekonomicznych nader częstym przypadkiem jest szacowanie pewnych wielkości, które są kombinacją dwóch różnych cech. Do takich wielkości może należeć:

1) przeciętna suma dochodów członków rodziny,
2) średnia różnica w wydatkach na określone dobro konsumpcyjne w dwóch typach zbiorowości,
3) wydajność plonów z hektara pól obsianych danym zbożem,
4) wzrost (lub spadek) dochodów danego typu firm w dwóch okresach.

a) Estymator sumy

Przy rozpatrywaniu więcej niż jednej cechy statystycznej może powstać problem estymacji sumy lub różnicy średnich obu cech. Rozróżnia się przy tym dwa przypadki:

Przypadek 1: Rozpatruje się jedną zbiorowość statystyczną o liczebności N w populacji generalnej i n w próbie, a warianty cech X oraz Y tworzą pary. Estymatorem sumy nieznanego parametru $\bar{X}+\bar{Y}$ w populacji generalnej jest

$$w_s=\bar{x}+\bar{y}. \tag{21}$$

Estymator (21) jest nie obciążony, a jego dyspersja wynosi

$$D^2(w_s)=D^2(\bar{x})+D^2(\bar{y})+2\rho_{xy}\sqrt{D^2(y)D^2(x)}=$$
$$=\left(\frac{N-n}{Nn}\right)(S_y^2+S_x^2+2\rho_{xy}S_yS_x). \tag{22}$$

b) Estymator różnicy

Analogicznie nieobciążonym estymatorem parametru $\bar{X}-\bar{Y}$ jest wyrażenie

$$w_d=\bar{x}-\bar{y}, \tag{23}$$

a jego dyspersja

$$D^2(w_d)=D^2(\bar{x})+D^2(\bar{y})-2\rho_{xy}D(\bar{x})\cdot D(\bar{y})=$$
$$=\left(\frac{N-n}{Nn}\right)S_x^2+S_y^2-2\rho_{xy}S_xS_y. \tag{24}$$

Wnioski:

1° Gdy $\rho_{xy}\approx 0$, wówczas $D^2(w_s)\approx D^2(w_d)$.

2° Dodatnia korelacja ($\rho_{xy}>0$) zwiększa wariancję sumy i pomniejsza wariancję różnicy.

3° Ujemna korelacja ($\rho_{xy}<0$) pomniejsza wariancję sumy i zwiększa wariancję różnicy.

Przypadek 2: Rozpatruje się dwie niekoniecznie jednakowo liczne próby wylosowane z populacji generalnej N elementów.

Warianty zmiennych X i Y nie tworzą w tym przypadku par. Nieobciążonym estymatorem różnicy jest

$$\hat{w}_d=\bar{x}-\bar{y}, \tag{25}$$

a dyspersja tego estymatora

$$D^2(\hat{w}_d)=\left(\frac{N-n_x}{N}\right)\frac{S_x^2}{n_x}+\left(\frac{N-n_y}{N}\right)\frac{S_y^2}{n_y}. \tag{26}$$

Jeśli próby n_x i n_y są tak małe w porównaniu z N, że poprawki na zwrotność losowania można nie brać pod uwagę, to

$$D^2(\hat{w}_d)=\frac{S_x^2}{n_x}+\frac{S_y^2}{n_y}. \tag{27}$$

Wzory na wariancję (26) i (27) odnoszą się także do sumy dwóch średnich w dwóch subpopulacjach.

c) Estymator ilorazu

Niech, jak poprzednio, R oznacza iloraz w populacji generalnej, a r – w próbie

$$R=\frac{\sum_{k=1}^{N} Y_k}{\sum_{k=1}^{N} X_k}=\frac{\bar{Y}}{\bar{X}}, \tag{28}$$

$$r=\frac{\sum_{i=1}^{n} y_i}{\sum_{i=1}^{n} x_i}=\frac{\bar{y}}{\bar{x}}. \tag{29}$$

Nadzieja matematyczna estymatora:

$$E\left(\frac{\bar{y}}{\bar{x}}\right)=\left(\frac{\bar{Y}}{\bar{x}}\right)\left[1+\left(\frac{N-n}{Nn}\right)(V_x^2-\rho_{xy}V_xV_y)\right], \tag{30}$$

gdzie

$$V_x=\frac{S_x}{\bar{X}}, \qquad V_y=\frac{S_y}{\bar{Y}},$$

ρ_{xy} – współczynnik korelacji między X i Y w populacji generalnej.

Jeśli zmienne Y_k i X_k występują w każdej dużej próbie losowej n, to dyspersja r wynosi w przybliżeniu:

$$D^2(r)=\frac{N-n}{NnX^2}\,\frac{\sum_{k=1}^{N}(Y_k-RX_k)^2}{N-1}. \tag{31}$$

3. Statystyka małych obszarów (SMO)

Wykorzystanie informacji o cesze dodatkowej leży też u podstaw statystyki małych obszarów (*small area statistics*).

Oprócz ogólnego pojęcia SMO używa się także określeń szczegółowych, takich jak: reprezentacyjne badanie małych obszarów (*small area sampling*), planowanie badania małych obszarów (*small area sample design*), estymacja parametrów dla małych obszarów (*small area estimation*).

Statystyka małych obszarów rozwinęła się w krajach, w których istnieje duże zapotrzebowanie na informację na niskim poziomie agregacji przestrzennej, jak na przykład powiaty i gminy. W wielu przypadkach także polskie województwa można zakwalifikować do małych obszarów.

Pole zastosowań SMO stale się poszerza, czemu sprzyja wzrost zapotrzebowania na informację w ujęciu regionalnym, lokalnym oraz o specyficznym charakterze (tzw. statystyka małych domen – *small domain statistics*), ekonomizacja badań oraz globalny przyrost informacji ogólnej. Pod względem przedmiotowym SMO znajduje zastosowanie najczęściej w przypadku rynku pracy i bezrobocia[1], zdrowia i zabezpieczenia społecznego, w biznesie (np. badanie rynku).

SMO nie należy utożsamiać z badaniami niewyczerpującymi. Istnieje znaczna grupa problemów związanych z pełną informacją, dotyczącą całej zbiorowości, która jednak z powodu swych rozmiarów sprawia, że dotyczące jej fakty ekonomiczne i społeczne są relatywnie mało liczne. Wielkość informacji zbieranych systematycznie (wyniki Narodowego Spisu Powszechnego, rejestracja bieżąca ruchu ludności itp.) przez Główny Urząd Statystyczny jest opracowywana w przekroju terytorialnym, według województw i dużych miast. Niektóre tego rodzaju dane są dostępne w podziale na gminy. Jednakże im niższy jest szczebel podziału terytorialnego kraju, tym mniejsza szczegółowość tablic statystycznych.

Jeśli więc nie chcemy stracić zbyt wiele na szczegółowości tablic statystycznych, a jednocześnie uniknąć większych błędów losowych, konieczne jest zastosowanie specjalnych procedur wypracowanych przez SMO. Przykładem takiego problemu może być projekcja demograficzna dla niewielkiego miasta, liczącego 50 - 100 tys. osób. Opracowanie takiej projekcji w sposób klasyczny, za pomocą metody składnikowej lub wieloregionalnej (wielostanowej) analizy demograficznej, pociągnie za sobą duże błędy projekcji. Dlatego warto zwrócić uwagę na paletę metod wypracowanych w ramach SMO.

Przy estymacji w SMO nawiążemy tylko do metod uwzględniających informacje o cesze dodatkowej, pomijając inne rodzaje estymatorów[2].

Jedno z możliwych podejść w SMO z wykorzystaniem dodatkowych informacji przedstawia C. Skinner [1991, s. 6 - 17].

---

[1] Por. np. J. D. Drew, M. P. Singh i G. H. Choudhry [1982], M. P. Singh i inni [1990], D. Elliot [1982].

[2] Inne estymatory syntetyczne opisali między innymi: G. H. Choudhry, J. N. K. Rao [1992], M. P. Singh, J. D. Gambino, H. J. Mantel [1992], W. L. Schaible [1992].

Wyróżnia trzy rodzaje zmiennych:

1) zmienne y odnoszące się do małego obszaru, odnotowane tylko w trakcie badania próbnego,

2) zmienne x, odnośnie do których informacje mamy zarówno z badania niewyczerpującego, jak i z innych źródeł, takich jak spis, rejestracja bieżąca,

3) zmienne z, które nie są obserwowalne w próbie, a informacje o nich można uzyskać na podstawie badania pełnego.

W SMO istnieje pewna klasa estymatorów nazywanych syntetycznymi (*synthetic estimators*), które uwzględniają w różny sposób dodatkowe informacje odnośnie do zmiennych typu x i z.

Należy też zwrócić uwagę na różnicę między estymatorem syntetycznym w SMO a wcześniej omawianymi klasami estymatorów: ilorazowych i liniowo-regresyjnych, gdzie zmienne typu z w ogóle nie występowały.

Przegląd estymatorów w SMO zaczniemy od formuły na konwencjonalny estymator regresyjny[3] (*conventional regression estimator*)

$$\bar{y}_{\mathrm{lr}\,a} = \bar{y}_a + \sum_{j=1}^{J} (\bar{X}_{aj} - \bar{x}_{aj})\beta_j, \tag{32}$$

gdzie: $\bar{x}_{aj}$ – estymator wartości $\bar{X}_{aj}$,

$j = 1, \ldots, J$ – numer podgrupy, warstwy itp.,

$\beta_1, \ldots, \beta_J$ – estymatory współczynników wielorakiej regresji liniowej otrzymane metodą najmniejszych kwadratów.

W klasie estymatorów syntetycznych możemy wyróżnić tzw. prosty syntetyczny estymator (*simple synthetic estimator*)

$$\bar{y}_{\mathrm{asyn}} = \sum_{j=1}^{J} w_{aj}\bar{y}_j, \tag{33}$$

gdzie w_{aj} to udział warstwy j w obszarze a.

Estymator (33) jest estymatorem obciążonym. Stosując go należy przyjąć założenie, że relacja między cechami Y, X jest taka sama w całym badanym obszarze.

Prosty syntetyczny estymator stanowi szczególny przypadek jednostkowego regresyjnego syntetycznego estymatora, który znajduje szerokie zastosowanie. Założenia dotyczące tego estymatora mogą być sprecyzowane później na podstawie modelu regresji, w pewnej mierze mogą one być także zmienione przy użyciu technik analizy regresji.

Jednostkowy syntetyczny estymator regresyjny ma następującą postać:

$$\bar{y}_{\mathrm{asyn\,reg}} = \sum_{j=1}^{J} \bar{X}_{aj}\beta_j, \tag{34}$$

---

[3] W opracowaniu C. Skinnera estymator ten nazywa się konwencjonalnym estymatorem regresyjnym.

gdzie: $\beta_1, \ldots, \beta_j$ – estymatory współczynników wielorakiej regresji liniowej,
$\bar{X}_{aj}$ – średnie w podobszarach j w regionie a.

W danym małym obszarze zmienna z przyjmuje tylko jedną wartość, dlatego nie może być wykorzystana w jednostkowej syntetycznej estymacji regresyjnej. Jeżeli jednak istnieje k zmiennych $z_1, \ldots, z_k$ ze znanymi średnimi $\bar{z}_{a1}, \ldots, \bar{z}_{ak}$ w obszarze a, to należy wziąć pod uwagę model regresji

$$y_a = \sum_{k=1}^{K} \bar{z}_{ak} \beta_{Bk} + \omega_a, \tag{35}$$

gdzie: $\beta_{B1}, \ldots, \beta_{BK}$ – współczynniki regresji, ze subskryptem B wskazującym na inny obszar,
ω_a – składnik losowy, którego średnia dla badanych obszarów równa jest zero.

Współczynniki regresji $\beta_{B1}, \ldots, \beta_{BK}$ można szacować na podstawie metody najmniejszych kwadratów.

Estymator $\bar{y}_a$ traktowany jest jako zmienna zależna, natomiast zmienna wartości $\bar{z}_{ak}$ jako zmienna niezależna. Pozwoli to utworzyć kolejny z estymatorów syntetycznych – regresyjny estymator syntetyczny na poziomie obszaru:

$$y_{\text{asynregarea}} = \sum_{k=1}^{K} \bar{z}_{ak} \beta_{Bk}. \tag{36}$$

Procedura ta została opisana przez Ericksena [1974] w celu szacowania zmian na określonym obszarze.

Literatura

Bracha, C. [1987], Wykorzystanie informacji o cechach dodatkowych w badaniach reprezentacyjnych, ZBSE GUS, PAN, Warszawa.

Bracha, C., Lednicki, B. [1989], Badanie efektywności różnych schematów losowania przy wykorzystaniu informacji o cechach dodatkowych, Wiadomości Statystyczne, nr 1.

Choudhry, G. H., Rao, J. N. K. [1992], Evaluation of small area estimators and empirical study, Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Cochran, W. G. [1976], Mietody wyborocznogo issledowanija, Statistika, Moskwa.

Drew, J. D., Singh, M. P., Choudhry, G. H. [1982], Evaluation of small area estimation techniques for the Canadian Labor Force Survey. Proceedings Survey Research Methodology, Section of the American Statistical Association.

Elliot, D. [1992], A project to examine the potential for small area estimates of employment and health data in the UK – Some preliminary thoughts, Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Jessen, R. J. [1985], Mietody statisticzeskich obsledowanij, Finansy i Statistika, Moskwa.

Platek, R., Rao, J. N. K., Särndal, C. E., Singh, M. P. [1987], Small area statistics. Invited presentations, Wiley, New York.

de Ree, J., van Huis, M. [1992], Synthetic estimation based on finite population models. Theory and application. Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Särndal, C. E. [1992], Design-based approaches in estimation for domains. Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Schaible, W. L. [1992], Use of small area estimators in U.S. Federal Programs, Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Singh, M. P., Drew, J. D., Gambino, J. G., Mayda, F. [1990], Methodology of the Canadian Labour Force Survey. Statistics Canada, Catalogue number 71 - 526.

Singh, M. P., Gambino, J. D., Mantel, H. J. [1992], Issues and options in the provision of small area data. Referat na Międzynarodową Konferencję Naukową, „Small Area Statistics and Survey Designs", GUS, Warszawa 30 września - 3 października 1992.

Skinner, C. [1991], The use of synthetic estimation techniques to produce small area estimates, Department of Social Statistics University of Southampton, maszynopis powielony.

Wywiał, J. [1992], Statystyczna metoda reprezentacyjna w badaniach ekonomicznych, Akademia Ekonomiczna im. Karola Adamickiego w Katowicach.

Zasępa, R. [1972], Metoda reprezentacyjna, PWE, Warszawa.

Zasępa, R. [1991], Zarys metody reprezentacyjnej, GUS, Warszawa.